ON THE ROLE OF MIXED STRATEGIES IN REPEATED GAMES WITH IMPERFECT PRIVATE MONITORING

Sekiguchi, Tadashi

(Citation)
Kobe University Economic Review,46:77-90

(Issue Date)
2000

(Resource Type)
departmental bulletin paper

(Version)
Version of Record

(JaLCDOI)
https://doi.org/10.24546/81000927

(URL)
https://hdl.handle.net/20.500.14094/81000927

ON THE ROLE OF MIXED STRATEGIES IN REPEATED GAMES WITH IMPERFECT PRIVATE MONITORING

By TADASHI SEKIGUCHI*

We consider repeated games with imperfect private monitoring, a class of repeated games that is most difficult to analyze. The main purpose is to reexamine the result by Sekiguchi [23], the first efficiency result in this context. The reexamination helps us understand how the idea of the result can be utilized in different situations. In particular, we show that the same type of efficiency result is obtained in a model of barter.

1. Introduction

In the field of repeated games in which seemingly every kind of Folk Theorem has been proved already, there is one class of those games that has eluded an attempt to establish a Folk Theorem. The class is called *repeated games with imperfect private monitoring,* and its primary feature lies in the assumption that each player observes a noisy signal about her opponents' past actions independently and privately, so that each player does not know precisely what the other players' signals are[1]. While our recognition is that we have a plethora of Folk Theorems, the truth is that all those results are obtained in a framework where the above assumption is *not* valid[2]. Namely, players do receive some (possibly imperfect) information about their rivals' past play, but it is assumed to be *public information.* We usually call this type of information (or monitoring) structure (imperfect) *public monitoring.*

Have the economists ignored the case of private monitoring and stayed on the realm of public monitoring just because private monitoring situations are of less economic

---

* The present paper is based on the talk the author gave at the Microeconomic Theory Conference, "Repeated Games with Private Monitoring", at the Cowles Foundation of Yale University. The author is thankful to the organizers of the conference for their hospitality, and thankful to the participants of the conference for their helpful comments and discussions.

1) Some authors call these games *repeated games with private monitoring*, omitting the term "imperfect". The reason for this omission is that imperfect information is necessary for monitoring to be private, so the term "imperfect" is redundant. Although this view is persuasive, I am also sympathetic to other authors' use of the term "imperfect private monitoring", because it has a clear purpose of distinguishing the situation from imperfect public monitoring, to be introduced later. For many years, the study of imperfect monitoring cases had concentrated on public monitoring, so that the term "imperfect monitoring" used to denote imperfect public monitoring. By explicitly using the term "imperfect" , we can alert the audience to the fact that we have one more important case of imperfect monitoring. Thus both terminologies are so appealing to me that I decide to use the two interchangeably throughout this paper.

2) It is a big enterprise to name only a few of those results: Aumann and Shapley [2] and Rubinstein [22] for nondiscounted payoffs, Friedman [10] and Fudenberg and Maskin [14] for the case of discounting, Benoit and Krishna [3] for finite repetition, Fudenberg, Levine and Maskin [13] for imperfect public information, and so on....

significance? The answer has to be "no". Let us consider, for instance, the "secret price-cutting" model by Stigler [24], in which firms have an opportunity to choose a lower price in secrecy. The only information a firm can get about the occurrence of such a secret price-cutting is its own sales level, which typically is its private information. Then the situation is private monitoring, to the extent that the sales level is affected not only by all the firms' prices but also by some unobservable shocks on the tastes of consumers. A more general and maybe more convincing example of private monitoring is observation errors. That is, players often recognize the other players' actions mistakenly. Since the other players have no way to know if a player commits such an observation error, the observed action is the private information of the player. Given that we can hardly expect that economic agents are too rational to make this type of mistakes, we can conclude that any economic model is subject to (possibly small) elements of private monitoring.

It seems to me that the only reason that much of the literature on repeated games has concentrated on public monitoring is that the private monitoring situation is very difficult to analyze[3]. One way to understand the difficulty is to invoke the result by Matsushima [17], who analyzes general repeated games with private monitoring, with the additional assumption that the privately observed signals are independent among players. Thus, given an action profile, knowing her own signal never helps a player learn the other players' signals better. Then Matsushima [17] shows that in general any pure strategy equilibrium of a repeated game with private monitoring must play a Nash equilibrium of the stage game in every period. In other words, we have to consider mixed strategies explicitly in order to obtain any positive result. In principle, however, we have little idea of what kind of mixed strategies will do. This apparent lack of a clue had been responsible for a very slow progress on repeated games with private monitoring.

Sekiguchi [23] is the first to prove a positive result for repeated games with private monitoring. In the context of repeated prisoners' dilemma, Sekiguchi [23] shows that the efficient (or cooperative) outcome can be approximated as an equilibrium if players are patient and if monitoring is sufficiently close to perfect monitoring. One appealing property of the equilibrium is the simplicity of its structure; it is a mixture of two very simple pure strategies. By the way, we should point out that the efficient result presented there is different from the standard Folk Theorem in the following ways. First, not every individually rational payoff vector is shown to be approximated. Second, while in standard Folk Theorems, like Fudenberg and Maskin [14] and Fudenberg, Levine and Maskin [13], the discount factor is the only variable that is adjusted in order to (approximately) achieve target payoffs, the efficiency result here adjusts both the discount factor *and* the monitoring structure. Namely, we first fix the payoff function, a relationship between a player's payoff and her action and signal, and then let monitoring structure converge to perfect monitoring.

---

3) By the way, an implicit assumption for this assessment is discounting. The case with no discounting or similar assumptions is studied by Radner [20] and Fudenberg and Levine [12]. We also assume communication among players is not allowed. With communication, strong results (including a Folk Theorem) are obtained by Compte [7] and Kandori and Matsushima [16]. See also Aoyagi [1].

Therefore we never fix a stage game, which is a big difference.

The purpose of this manuscript is to reexamine the result of Sekiguchi [23] and to explore the possibility that the same type of mixed strategies enables us to prove a similar efficiency result in different contexts. This project is of particular interest when we look at developments of recent research on repeated games with private monitoring. We have some papers that prove a similar efficiency result for prisoners' dilemma and some similar games, using a quite different type of mixed strategies. Among those are Piccione [19] and Ely and Valimaki [9] (See also Matsushima [18] as a related work), and their equilibrium strategies are such that players randomize in almost all periods. Thus we have two philosophies about how to derive a positive result, and it is important to understand what one approach can do and what the other can do. Here we attempt to understand what the approach by Sekiguchi [23] can do.

The reexamination of the result by Sekiguchi [23] is also the subject of Bhaskar and Obara [6]. Their main result is to extend the efficiency result to more general prisoners' dilemma, including its n-player version. This manuscript extends it to a different situation. Namely, we consider a model of barter, played by n traders. The model has a similar structure to prisoners' dilemma, in the sense that while it is efficient for the traders to exert efforts before the goods are exchanged, they have an incentive to avoid the costly efforts. If the realized quality of the goods is only a noisy signal of the traders' efforts, then the situation is private monitoring because it is reasonable to assume that the quality of the goods is private information of the consumer. The main result from analysis of this model is that the efficient outcome can be approximated by an equilibrium, using the same idea as Sekiguchi [23].

The rest of the paper is organized as follows. Section 2 introduces the basic model. Section 3 provides a fundamental way, or a "cookbook", to find a candidate equilibrium strategy and to examine its equilibrium property. Section 4 describes our model of barter in detail, and then shows how the efficiency result is obtained by simply applying the cookbook. Section 5 presents the conclusions.

2. The Model

Despite our restricting attention to more specific cases in later sections, we attempt to develop a more general description of the model in this section. The reason we stick to the generalization is twofold. First, having a general model facilitates description of more specific models in later sections. Second, the expression developed here contains some useful notions, to be explored in future research.

We begin with the stage game, denoted by G, played by n players who simultaneously choose their actions. Let A_i be player i's set of actions, which is assumed to be finite. Each player can play a mixed strategy, so let S_i be the set of mixed actions of player i. Let $A = \times_{i=1}^{n} A_i$ and $S = \times_{i=1}^{n} S_i$ be the set of pure and mixed action profiles, respectively. After choosing actions, each player i observes a signal about $a \in A$, which belongs to a finite set Ω_i. One important assumption is that player $j \neq i$ cannot observe ω_i. We define $\Omega = \times_{i=1}^{n}$

Ω_i, which is the set of signal profiles. The signals are stochastic and we denote the probability that signal profile $\omega \in \Omega$ is realized under action profile $a \in A$ by $\pi(\omega|a)$. Each player's payoff depends only on her action and signal, denoted by $u_i(a_i, \omega_i)$ for player i [4]. Thus G is determined by n (the number of players), A (the set of pure action profiles), Ω (the set of signal profiles),$\{u_i\}_{i=1}^n$ (payoff functions) and π (monitoring structure).

Given $G = \{n, A, \Omega, \{u_i\}_{i=1}^n, \pi\}$, let us define the normal-form game with the same number of players and the same action set as G, whose payoff function of player i is defined by

$$g_i(a) = \sum_{\omega \in \Omega} \pi(\omega|a)\, u_i(a_i, \omega_i).$$

We call this game, $\Gamma = \{n, A, \{g_i\}_{i=1}^n\}$, the *normal-form representation of G*.

Let G and its normal-form representation Γ be given. $G^\infty(\delta)$ is the game in which G is played infinitely often, starting from period 1, with discount factor δ. In this repeated game, player i's *history* at the beginning of period $t \geq 2$ (we often call it simply a history at period t) consists of her past actions and private signals, denoted by $h_i^t = \{(a_i(\tau), \omega_i(\tau))\}_{\tau=1}^{t-1}$. Let H_i^t ($t \geq 2$) be the set of histories of player i at period t. Defining H_i^1 as an arbitrary singleton, we define a (behavioral) strategy of player i in $G^\infty(\delta)$ as a mapping from $\bigcup_{t=1}^\infty H_i^t$ to S_i. Let Σ_i be the set of strategies of player i in $G^\infty(\delta)$, and let us define $\Sigma = \times_{i=1}^n \Sigma_i$. The overall payoffs of player i is the average discounted sum of the payoffs in each period. Therefore, strategy profile $\sigma \in \Sigma$ gives player i the payoffs of

$$f_i(\sigma) = (1-\delta)E\left[\sum_{t=1}^{\infty} \delta^{t-1} g_i(a(t))\right],$$

where expectation is taken given σ and π, and $a(t)$ is the action profile played in period t.

We shall consider the case in which monitoring is noisy in the sense that any signal can be observed under any action profile, but we shall also assume that the probability that a wrong signal is observed given an action profile is small. This motivates the following assumptions and terminology.

Assumption 1 *For any* $\omega \in \Omega$ *and* $a \in A$, $\pi(\omega | a) > 0$.

Assumption 2 $\Omega_i = A_{-i} = \times_{j \neq i} A_j$ *for any* i.

Definition. Suppose Assumption 2 holds. π is *ε-perturbed* if for any i and any $a \in A$, we have

$$\pi_i(a_{-i}|a) \geq 1 - \varepsilon, \tag{1}$$

where $\pi_i(\omega_i|a)$ is the marginal distribution of $\pi(\omega|a)$.

Assumption 1 is the standard full support condition. Note that Assumption 2 makes the expression $\pi_i(a_{-i}|a)$ in (1) relevant, which denotes the probability that the signal a_{-i}, as an

4) Note that this assumption is natural for the secret price-cutting model we have seen before: the action corresponds to the actual price, and the signal to the sales level. Those variables solely determine the firm's profit.

element of Ω_i, realizes given a. This a_{-i} should not be confused with the actions played by all the players other than i. One way to justify Assumption 2 is to assume that the source of private monitoring in this model is observation errors. Namely, there is a possibility that each player perceives that her contenders play a_{-i}, despite the fact that their actions are $a'_{-i} \neq a_{-i}$. This interpretation also motivates the notion of ε-perturbation, because ε-perturbed monitoring for small $\varepsilon > 0$ captures the idea that players are subject to observation errors but they are not so stupid as to make such mistakes that often. By the way, note that if monitoring technology is ε-perturbed, it is ε'-perturbed for any $\varepsilon' > \varepsilon$.

Given a normal-form game $\Gamma = \{n, A, \{g_i\}_{i=1}^n\}$, we define $G(\Gamma)$ as the set of all $G = \{n, A, \Omega, \{u_i\}_{i=1}^n, \pi\}$ satisfying Assumptions 1 and 2 whose normal-form representation is Γ. Similarly, we define $G(\Gamma, \varepsilon)$ as the set of all $G \in G(\Gamma)$ with ε-perturbed π.

3. A Cookbook

In this section, we discuss a candidate for an (possibly efficient or cooperative) equilibrium, and how to check the equilibrium property of the candidate strategy profile, by giving a detailed account of the result by Sekiguchi [23]. As a result, we will have a cookbook of the situation. During the process, we will also understand the idea of subsequent work like Bhaskar and Obara [6].

Let us specify the model used throughout in this section, which is a repeated prisoners' dilemma. We first specify Γ, which is given by $n = 2$, $A_1 = A_2 = \{C, D\}$, and $g_1(C, C) = g_2(C, C) = 1$, $g_1(D, C) = g_2(C, D) = 1 + \alpha$, $g_1(C, D) = g_2(D,C) = -\beta$ and $g_1(D, D) = g_2(D, D) = 0$, where $\alpha > 0$, $\beta > 0$ and $\alpha - \beta < 1$. We also need an additional assumption of symmetric monitoring.

Assumption 3 *For any $a \in A$ and $\omega \in \Omega$,*

$$\pi(\omega|a) = \pi((\omega_2, \omega_1)|(a_2, a_1)). \tag{2}$$

Given Γ, let $G^*(\Gamma, \varepsilon)$ be the set of all elements of $G(\Gamma, \varepsilon)$ with π satisfying Assumption 3.

In the following analysis, two important ideas play a central role. First, although we know that sequential equilibrium is more compelling a solution concept than Nash equilibrium, *we deliberately start our analysis by finding a nearly efficient Nash equilibrium*. Indeed, the full support assumption (Assumption 1) justifies our attitude, because any Nash equilibrium has an outcome-equivalent sequential equilibrium if this assumption is satisfied.[5] Later, we will see that working with Nash equilibrium simplifies the analysis because we can work with a simple strategy profile that is unlikely to be a sequential equilibrium.

Second, despite our main concern about patient players, *we first consider the case with players who are not so patient.* We can do so because of some monotonicity results of the equilibrium payoff set with respect to δ. Indeed, if public randomization is available, the

5) See Sekiguchi [23] or Kandori and Matsushima [16] for a formal proof.

monotonicity is immediate; see Bhaskar and van Damme [5]. Even if public randomization is *not* available, we can use the monotonicity result by Ellison [8]. Given those results, we can concentrate on the case with relatively heavy discounting.

Now we can proceed to the details of the cookbook, which consists of three steps.

Step 1. Find a candidate strategy. As Matsushima [17] suggests, we surely have to consider mixed strategies because our model includes the case of conditional independence.[6] On the other hand, we want to have as simple a strategy as possible. To satisfy both demands at the same time, Sekiguchi [23] proposes the following generalization of the grim trigger strategies of G^{∞}. First, the grim trigger strategy, denoted by σ_C, is defined as

$$\sigma_C(h_i^t) = \begin{cases} C & \text{if } t = 1 \text{ or if } h_i^t = \{(C,C),(C,C),\cdots,(C,C)\} \\ D & \text{otherwise} \end{cases}$$

Next, let σ_D be the strategy such that $\sigma_D(h_i^t) = D$ at any h_i^t.

Let us define σ^* as the strategy in which a player randomizes between σ_C and σ_D so that the opponent is indifferent between σ_C and σ_D. Note that symmetry of stage game payoffs and monitoring (the latter is captured by (2)) guarantee that σ^* is the same for each player if it exists. Let q^* be the probability of σ_D attached by σ^*. We often write the strategy in which σ_D is chosen with probability λ and σ_C with probability $1-\lambda$ as $\lambda\sigma_D + (1-\lambda)\sigma_C$. Thus we have $\sigma^* = q^*\sigma_D + (1-q^*)\sigma_C$.

A simple computation shows that if monitoring is ε-perturbed for sufficiently small $\varepsilon > 0$ and if $\delta > \frac{\alpha}{1+\alpha}$, σ^* exists. More importantly, σ^* approximates the payoff of cooperation (namely, 1) if monitoring is ε-perturbed for sufficiently small $\varepsilon > 0$ and if δ is sufficiently close to $\frac{\alpha}{1+\alpha}$. This is because the corresponding q^* is close to zero, and cooperation would not collapse easily.

This profile (σ^*, σ^*) is our equilibrium candidate. By the above argument, the efficiency result is established immediately if we show that (σ^*, σ^*) is a Nash equilibrium whenever π is ε-perturbed for any sufficiently small ε and δ is sufficiently close to $\frac{\alpha}{1+\alpha}$.

Step 2. Check sequential rationality. The next step is to understand what continuation strategy is optimal at any possible history; in other words, to check sequential rationality. The task is actually very difficult to do, because the system of beliefs of a player tends to be highly complicated in private monitoring situations.

However, the candidate strategy and our line of attack make this task less demanding. First, note that our candidate strategy σ^* is such that the continuation strategy at any history has the form of $\lambda\sigma_D + (1-\lambda)\sigma_C$. Thus a player's belief about the opponent's continuation strategy at some history is summarized into a one dimensional variable λ. Second, since we

---

6) Namely, for any $\omega \in \Omega$ and $a \in A$,

$$\pi(\omega | a) = \prod_{i=1}^{n} \pi_i(\omega_i | a)$$

are interested in the Nash equilibrium property of the profile, all we have to consider is the histories on the path.

Before we proceed, however, we have to understand what happens in the case of *perfect monitoring*. Fix Γ, and consider a repeated game where Γ is played infinitely often under the assumption of perfect monitoring, with discount factor δ. We denote this game by $\Gamma^{\infty}(\delta)$. We can define σ^* for $\Gamma^{\infty}(\delta)$, too. If $\delta > \frac{\alpha}{1+\alpha}$, σ^* exists with $q^* > 0$. Let us define Σ_i^C (Σ_i^D, respectively) as the set of strategies of player i that assigns $C(D)$ in the first period. Now the following observations are immediate. First, σ_C is strictly better than any $\sigma_i \in \Sigma_i^D$ if player j plays $\lambda\sigma_D + (1-\lambda)\sigma_C$, where $\lambda < q^*$. Second, σ_D is strictly better than any $\sigma_i \in \Sigma_i^C$ if player j plays $\lambda\sigma_D + (1-\lambda)\sigma_C$, where $\lambda > q^*$. The following result tells us that these observations can be extended to private monitoring situations, provided that monitoring is almost perfect.

Proposition 1 *Let Γ and $\delta > \frac{\alpha}{1+\alpha}$ be given. Let q^* be the probability of σ_D in σ^*, defined for $\Gamma^{\infty}(\delta)$. Then for any $\eta > 0$, there exists $\varepsilon > 0$ such that for any $G \in G^*(\Gamma, \varepsilon)$, $G^{\infty}(\delta)$ satisfies the following.*

1. *It is not sequentially rational for player i to play $\sigma_i \in \Sigma_i^D$ if player j plays $\lambda\sigma_D + (1-\lambda)\sigma_C$, where $\lambda < q^* - \eta$.*
2. *It is not sequentially rational for player i to play $\sigma_i \in \Sigma_i^C$ if player j plays $\lambda\sigma_D + (1-\lambda)\sigma_C$, where $\lambda > q^* + \eta$.*

Proof. We first claim that we can endow Σ with a topology with respect to which (1) for fixed π (whether perfect or private monitoring), the payoff function $f_1(\sigma)$ and $f_2(\sigma)$ are a continuous function of σ, and (2) Σ is compact.

Note that each S_i is identified with a finite dimensional Euclidean space, so that we can topolize it with the Euclidean distance. Since each Σ_i is a product space of countably many S_i's, we can define a topology of Σ_i as the product topology of the S_i's. Then we can define a topology of Σ as the product topology of Σ_1 and Σ_2. Applying the argument by Fudenberg and Levine [11], it is easy to prove that $f_1(\sigma)$ and $f_2(\sigma)$ are continuous in σ, according to this topology.[7)] Compactness of Σ follows from Tychonov's theorem (see Royden [21]).

Note that if monitoring is perfect, we have

$$f_1(\sigma_C, \lambda\sigma_D + (1-\lambda)\sigma_C) > \max_{\sigma_1 \in \Sigma_1^D} f_1(\sigma_1, \lambda\sigma_D + (1-\lambda)\sigma_C) \tag{3}$$

for any $\lambda < q^*$. Since Tychonov's theorem also implies that each Σ_i is compact with respect to the product topology, so is Σ_i^D. Thus the above argument implies that the RHS of (3) is a uniformly continuous function of λ, by the Theorem of the maximum. Since it is easily seen that $f_1(\sigma)$ is continuous with respect to monitoring technology (note that monitoring

7) See also Fudenberg and Tirole [15] for a proof.

structure can be identified with a finite dimensional Euclidean space), the first part of the proposition follows for player 1. The case of player 2 and the second part can be done in a similar way.

Therefore, we have almost full characterization about the best response given a belief. Although a neighborhood of q^* is not covered by this characterization, we will see that we do not have to worry about that, because the belief never falls on the range on the path given σ^*, except the very first period of the game.

Step 3. Check the dynamics of beliefs. Let us start with specifying what we want to do in this step. First, fix Γ, and choose $\delta > \frac{\alpha}{1+\alpha}$ so that it is close to $\frac{\alpha}{1+\alpha}$, and therefore q^*, defined as in Proposition 1, is close to zero. Next, choose $\eta > 0$ so small that both $q^* - \eta > 0$ and $q^* + \eta < 1/2$ hold. Let $\varepsilon > 0$ be the corresponding value given in Proposition 1. Now we want to show that there exists $\varepsilon_1 \in [0, \varepsilon]$ such that for any $G^\infty(\delta)$ with $G \in G(\Gamma, \varepsilon_1)$, the corresponding σ^* satisfies the following properties.

(1) Given the profile (σ^*, σ^*), at any h_i^t on the path that assigns C, where $t \geq 2$, player i believes that player j plays σ_D with a probability smaller than $q^* - \eta$.
(2) Given the profile (σ^*, σ^*), at any h_i^t on the path that assigns D, where $t \geq 2$, player i believes that player j plays σ_D with a probability greater than $q^* + \eta$.

In view of Proposition 1, these properties guarantee that the profile (σ^*, σ^*) is such that each player has no incentive to deviate at any history on the path except the initial history. However, since each player is indifferent between σ_C and σ_D at the initial history by the very definition of σ^*, no profitable deviation exists at the initial history. Thus (σ^*, σ^*) is a Nash equilibrium.

Let us check whether we can find such ε_1. First, note that if monitoring is ε'-perturbed for ε' much smaller than q^*, the signal in the first period is very informative about the opponent's choice between σ_C and σ_D. In particular, if player i observes C (D, respectively) in the first period, then she would assume that j selected C (D) in the first period and therefore his continuation strategy is quite likely to be σ_C (σ_D), irrespective of her action in period 1. Hence the above properties are satisfied for any history at period 2, if monitoring is sufficiently close to perfect monitoring.

Next, we consider the histories at period $t \geq 3$ which is on the path and assigns C. Those histories must have the form of $h_i^t = \{(C,C),(C,C),\cdots,(C,C)\}$. If monitoring is almost perfect, this history is a signal that player j started with σ_C and still remains cooperative. Moreover, we can show that almost perfect monitoring implies that player i at $h_i^t = \{(C,C),(C,C),\cdots,(C,C)\}$ is more convinced that player j's continuation strategy is σ_C than she was at $h_i^{t-1} = \{(C,C),(C,C),\cdots,(C,C)\}$. Thus, if monitoring is so close to perfect that the property (1) is valid at $h_i^3 = \{(C,C),(C,C)\}$, then it is also valid at all the histories that assigns

C. Thus the property (1) is satisfied.

Finally, in order to establish the property (2), we consider the histories at period $t \geq 3$ which is on the path and assigns D. We have two kinds of such histories. First, let us consider h_i^t satisfying $a_i(t-1) = C$ and $\omega_i(t-1) = D$. Since h_i^t is on the path, we must have $h_i^t = \{(C,C),(C,C),\cdots,(C,C),(C,D)\}$. For the opponent to be still cooperative, $h_j^t = \{(C,C),(C,C),\cdots,(C,C)\}$ must be the case. Thus the D observed in period t-1 is an observation error, and it is the only observation error observed by any player in the course of play. Now let us consider the following history; $h_j^t = \{(C,C),(C,C),\cdots,(C,C),(C,D),(D,C)\}$ at which player j's continuation strategy is σ_D. If player j is at this history, then the D observed by player j in period t–2 (recall $t \geq 3$) is the only observation error in the course of play; the D observed by player i now reflects the fact that j had switched to σ_D correctly. Because of symmetry of monitoring (namely, (2)), player i must believe that the above two histories of player j are equally likely. Since player j's other histories require him to play σ_D, player i believes that player j plays σ_D with a probability greater than 1/2. Note that this argument applies to any history with the form $h_i^t = \{(C,C),(C,C),\cdots,(C,C),(C,D)\}$ if monitoring is almost perfect. Since $q^* + \eta < 1/2$, the property (2) is established.

Finally, we consider the type of histories h_i^t, where $a_i(t-1) = D$. Since h_i^t is on the path, there must exist $T \geq 1$ such that $a_i(\tau) = C$ for any $\tau < T$ and $a_i(\tau) = D$ for any $\tau \geq T$. We first consider the case when $T = t-1$. Then $h_i^t = \{(C,C),(C,C),\cdots,(C,C),(C,C),(D,\omega_i(t-1))\}$ must hold. The above argument shows that player i was convinced that the opponent was already defective with a probability no smaller than 1/2 at the beginning of period $t-1$. Thus, if $\omega_i(t-1) = D$, this observation simply confirms her suspicion that the opponent has already switched to σ_D. If $\omega_i(t-1) = C$, then almost perfect monitoring ensures that player j still playing according to σ_C would observe D in period $t-1$ with a probability close to 1 and therefore switch to σ_D from now on. In any case, player i is still convinced that the opponent was already defective with a probability no smaller than 1/2, given h_i^t. Note that this argument applies to any history with $T = t-1$, if monitoring is sufficiently close to perfect. The case with a different value of T is treated inductively, and we can see that the result follows for any h_i^t with $a_i(t-1) = D$, if monitoring is sufficiently close to perfect.

Combining all those arguments, the properties (1) and (2) hold if monitoring is almost perfect. Thus we have shown that efficient outcomes can be approximated as a Nash equilibrium for some discount factors. As has been discussed before, the extension to sequential equilibrium and patient players is straightforward. Therefore, we finally obtain the efficiency result.

Note that the above argument has nothing to do with the payoff structure of the model, as far as it corresponds to a prisoners' dilemma. Thus the restriction on payoffs assumed by Sekiguchi [23] is unnecessary. In fact, Sekiguchi [23] needs the restriction because the argument is a bit sloppy about characterization of best responses (Step 2 in our cookbook). The characterization is much more thoroughly done by Bhaskar and Obara [6], who first show that the idea of Sekiguchi [23] still applies if the restriction is removed, but the main feature of their argument is captured by our argument of Step 2.

4. A Model of Barter

In this section, we consider how the cookbook described in the previous section can be applied to a different situation. To this end, we present the following model of barter.

We start with description of the stage game. We have $n \geq 3$ players, or traders, each of whom possesses goods to be traded. Each trader i is interested in consuming the goods which is initially at the hands of player $i + 1$. Throughout the analysis, we adopt the convention that $n + 1 = 1$. They are not interested in their own goods at all, so each player i will automatically give the goods initially held by herself to player $i - 1$ (the convention is $1 - 1 = n$). Before doing so, however, they have two actions to choose from; to make a proper effort before giving the goods (C), or not to make such an effort (D). Whether they make an effort or not, delivery of those goods is simultaneous, so the timing of choice of efforts can also be considered as simultaneous. Thus we have $A_i = \{C,D\}$ as the set of pure actions of player i.

If player $i + 1$ selects C, the quality of the goods consumed by i is good with a large probability. However, despite the proper effort, the quality of the goods can be bad, as if no efforts had been made. We also assume that with some small probability, no efforts produce good quality. We allow the good quality and the bad quality only. Hence Ω_i has only two elements, which is a departure from Assumption 2 (recall that $n \geq 3$). Since the actual quality of the goods is private information of the consumer, we have private monitoring. Formally, we assume the following on the monitoring structure, in addition to Assumption 1 (full support).

Assumption 4 *For any i, we have $\Omega_i = \{C,D\}$.*

Assumption 5 *For any $a \in A$ and $\omega \in \Omega$, we have*

$$\pi_i(\omega_i|a) = \pi_i(\omega_i|(a_i, a_{i+1})), \tag{4}$$

and $\pi_i(\omega_i|(a_i, a_{i+1}))$ *does not depend on i. We also have*

$$\pi(\omega|a) = \prod_{i=1}^{n} \pi_i(\omega_i|(a_i, a_{i+1})). \tag{5}$$

Those assumptions imply that the quality of the goods player i receives (stochastically) depends solely on (a_i, a_{i+1}), the dependence is independent of the choice of pair $(i, i + 1)$(by (4)), and those qualities are independent among the traders, given an action profile (by (5)).

As for payoff functions, we begin with the payoff functions of the corresponding normal-form representation of the stage game. For any $a \in A$, we assume that player i's payoff is determined solely by (a_i, a_{i+1}). In particular, it is assumed that

$$g_i(a) = g_i(a_i, a_{i+1}) = \begin{cases} 1 & \text{if } a_i = a_{i+1} = C \\ 1+\alpha & \text{if } a_i = D \text{ and } a_{i+1} = C \\ -\beta & \text{if } a_i = C \text{ and } a_{i+1} = D \\ 0 & \text{if } a_i = a_{i+1} = D \end{cases} \tag{6}$$

where $\alpha > 0$, $\beta > 0$ and $\alpha - \beta < 1$. Note that the same α and β apply to any player. More importantly, note also that each player has no incentive to make an effort if the game is one-shot, which makes it similar to prisoners' dilemma.

For $\Gamma = \{n, A, \{g_i\}_{i=1}^{n}\}$, where each g_i is given by (6), let $G^{**}(\Gamma)$ be the set of all G satisfying Assumptions 1, 4 and 5 whose normal-form representation is Γ. Let us also define $G^{**}(\Gamma, \varepsilon)$ as the set of all elements of $G^{**}(\Gamma)$ which also satisfies that;

$$\pi_i(a_{i+1}|(a_i, a_{i+1})) \geq 1 - \varepsilon \tag{7}$$

for any i, $a_i \in A_i$ and $a_{i+1} \in A_i$. With an abuse of terminology, we call π satisfying (7) ε-perturbed. ε-perturbation for small ε implies that the payoff function of player i of G is close to g_i given by (6). Thus (6) reflects plausible assumptions that efforts are costly and that traders like good quality.

Fix Γ, and consider $G^{\infty}(\delta)$ for some $G \in G^{**}(\Gamma)$ and δ. We examine whether the efficient outcome, which is the action profile $(C, C, \cdots, C)$ by (6), can be approximated as an equilibrium, using the cookbook we developed before.

Step 1. Find a candidate strategy. This step is not difficult. Note that in $G^{\infty}(\delta)$, each player chooses either C or D and then observes C or D in each period. This structure is the same as the repeated prisoners' dilemma considered in the previous section. Thus σ_C and σ_D are defined in the same way, as well as σ^* and q^*. We also define Σ_i^C and Σ_i^D in the same way. The profile $(\sigma^*, \sigma^*, \cdots, \sigma^*)$ is our equilibrium candidate. Note, however, that existence of the profile now requires ε-perturbation for small ε and $\delta^{n-1} > \frac{\alpha}{1+\alpha}$. The profile $(\sigma^*, \sigma^*, \cdots, \sigma^*)$ is nearly efficient if π is ε-perturbed for small $\varepsilon > 0$ and if δ is sufficiently close to $(\frac{\alpha}{1+\alpha})^{\frac{1}{n-1}}$.

Step 2. Check sequential rationality. As we did in Section 3, we start from the case of *no noise*, the case in which each player i always correctly observes player $(i+1)$'s action. Note that no noise is not equivalent to perfect monitoring, because player i cannot observe the action chosen by any player $j \neq i+1$.

Given Γ, let us denote the infinitely repeated game with no noise by $\Gamma^{\infty}(\delta)$, where δ is the discount factor. Suppose player i believes that any player $j \neq i$ plays either σ_C or σ_D. Let us summarize her belief in the following *belief vector* $\mu = (\mu_1, \mu_2, \cdots, \mu_{n-1})$, where μ_k denotes the probability that player $i+k$ chooses σ_D and that all the players $i+1, i+2, \cdots, i+k-1$ choose σ_C. Thus $\Sigma_{k=1}^{n-1} \mu_k$ denotes the probability that at least one player plays σ_D. Note that given the profile $(\sigma^*, \sigma^*, \cdots, \sigma^*)$, player i's belief vector at the initial history, μ^*, is such that $\mu_k^* = q^*(1-q^*)^{k-1}$.

Now we have the following results for $\Gamma^{\infty}(\delta)$. First, if $\mu_k = 1$ for some k, σ_D is strictly better than any $\sigma_i \in \Sigma_i^C$. This is because player i cannot change when player $i+1$ switches to σ_D (note that this anyway occurs) by changing her own play. Thus her current action has no effects on what happens in the future, which makes σ_D optimal. Playing $\sigma_i \in \Sigma_i^C$ just results in losing money. Second, if $\mu_1 > (n-1)\,q^*$ and $\mu_k = 0$ for any other k, σ_D is strictly better than any $\sigma_i \in \Sigma_i^C$. To see this, note that given $(\sigma^*, \sigma^*, \cdots, \sigma^*)$, the probability that at least one of the $n-1$ players chooses σ_D is less than $(n-1)\,q^*$. Thus, if player $i+1$ is defective with a larger probability, σ_D is strictly better than σ_C for player i, even if the other players coordinate on σ_C with the remaining probability. Since it is easily seen that σ_C is best on Σ_i^C given the belief, the claim follows. Third, if $\mu_k = 0$ for some k and $\mu_j = \mu_j^*$ for any other j, σ_C is strictly better than any $\sigma_i \in \Sigma_i^D$. This is because the definition of σ^* implies that σ_C is strictly better than σ_D. Since the first result shows that σ_D is best on Σ_i^D, the claim follows.

These results state that in several circumstances either σ_C or σ_D is a strictly better reply than any element of Σ_i^D or Σ_i^C. Thus the continuity argument that we used in Proposition 1 shows the following result, whose proof is omitted.

Proposition 2 *Let Γ and $\delta > (\frac{\alpha}{1+\alpha})^{\frac{1}{n-1}}$ be given. Let q^* be the probability of σ_D in σ^*, defined for $\Gamma^{\infty}(\delta)$. Then there exist $\eta > 0$ and $\varepsilon > 0$ such that for any $G \in G^{**}(\Gamma, \varepsilon)$, $G^{\infty}(\delta)$ satisfies the following. Suppose player i believes that any player $j \neq i$ play either σ_C or σ_D, and let μ be the corresponding belief vector. Then*

(1) If $\mu_k \geq 1 - \eta$ for some k, σ_D is strictly better than any $\sigma_i \in \Sigma_i^C$.
(2) If $\mu_1 > (n-1)q^*$ and $\mu_k = 0$ for any other k, σ_D is strictly better than any $\sigma_i \in \Sigma_i^C$.
(3) If $\mu_k \leq \eta$ for some k and $\mu_j \leq \mu_j^* + \eta$ for any other j, σ_C is strictly better than any $\sigma_i \in \Sigma_i^D$.

Step 3. Check the dynamics of beliefs. Now we fix Γ and $\delta > (\frac{\alpha}{1+\alpha})^{\frac{1}{n-1}}$. We choose δ so close to $(\frac{\alpha}{1+\alpha})^{\frac{1}{n-1}}$ that q^* corresponding to σ^*, defined for $\Gamma^{\infty}(\delta)$, is such that $(n-1)q^* < 1/2$. Let $\eta > 0$ and $\varepsilon > 0$ be the corresponding values in Proposition 2. Our task is to find $\varepsilon_1 \in [0, \varepsilon]$ such that for any $G^{\infty}(\delta)$, where $G \in G^{**}(\Gamma, \varepsilon_1)$, σ^* defined for $G^{\infty}(\delta)$ is such that the profile $(\sigma^*, \sigma^*, \cdots, \sigma^*)$ is a Nash equilibrium. To this end, we shall use Proposition 1 and show that any history on the path at period $t \geq 2$ has the belief vector that makes the action attached to the history optimal. If this is the case, $(\sigma^*, \sigma^*, \cdots, \sigma^*)$ is Nash equilibrium because we have no profitable deviation at period 1, thanks to the definition of σ^*.

First, consider the histories on the path that assign C. Thus it follows that $h_i^t = \{(C,C),(C,C),\cdots,(C,C)\}$. Suppose $t = 2$. Then no noise would imply that the corresponding belief vector satisfies $\mu_k = \mu_k^*$ for any $k \leq n - 2$, and $\mu_{n-1} = 0$. Hence, if monitoring is nearly perfect, the belief vector would satisfy the condition of Proposition 2(3). Thus playing C is optimal. The same type of argument applies when $t \geq 3$.

Second, consider the histories on the path that assign D. The case $h_i^t = \{\cdots,(D,\cdot)\}$(including the case $t = 2$) is easiest, because ε-perturbation, where $\varepsilon \leq \eta$, ensures that player $i-1$ observed D in period $t-1$ and his continuation strategy is σ_D with a probability greater than $1 - \eta$. Thus Proposition 2(1) guarantees optimality of playing D. If $h_i^2 = \{(C,D)\}$, ε-perturbation with ε being small compared to q^* implies that player $i+1$ selected σ_D with a probability close to one. Thus optimality of playing D easily follows. The remaining case is $h_i^t = \{(C,C),(C,C),\cdots,(C,C),(C,D)\}$, where $t \geq 3$.[8] For the other players to play σ_C, it must be that $h_j^t = \{(C,C),(C,C),\cdots,(C,C)\}$ for any $j \neq i$, which implies that the D observed in period $t-1$ is an error. Given h_i^t, let us consider the event that $h_j^t = \{(C,C)(C,C),\cdots,(C,C)\}$ for any $j \neq i+1$ and $h_{i+1}^t = \{(C,C),(C,C),\cdots,(C,C)(C,D),(D,C)\}$. This event also has a single error; the D in period $t-2$ player $i+1$ observed. Due to the symmetry assumption (Assumption 5), those two events are equally likely. Thus, conditional that one of the two events occurs, player i's (conditional) belief vector is such that $\mu_1 = 1/2$ and $\mu_k = 0$ for any other k. Since any other possible event entails at least one player playing σ_D, the belief vector given h_i^t is a convex combination of $(1/2,0,0,\cdots,0)$ and belief vectors such that $\mu_k = 1$ for some k. Therefore, for each of the belief vectors that form the convex combination, Proposition 2(1) and (2) show that σ_D dominates any $\sigma_i \in \Sigma_i^C$. Therefore, for the belief vector at h_i^t, σ_D also dominates any $\sigma_i \in \Sigma_i^C$, which proves optimality of playing D.

Combining all the arguments, we prove that $(\sigma^*, \sigma^*,\cdots, \sigma^*)$ is a Nash equilibrium, if monitoring is almost perfect. The same argument as in Section 3 extends the result to sequential equilibrium and to patient players.

5. Conclusions

The present paper examines how the idea of using relatively simple mixed strategies is helpful in proving an efficiency result in repeated games with private monitoring, based on the result by Sekiguchi [23]. The main conclusion is that we can apply the idea successfully to a model of barter.

One might want to argue that the extension made here is not very distant from where we were, even if we include the extension by Bhaskar and Obara [6]. From this viewpoint, the idea proposed by Sekiguchi [23] might be of less significance. However, the constructed equilibrium possesses a very appealing feature of simplicity. If we have a model that can be solved both by the idea presented here and by the approach of Piccione [19] and others, we might be tempted to think that the equilibrium supported by the former idea is easier to play.[9] Therefore, we believe that it is still important to examine when such simple strategies

8) Any other history would not be on the path.

9) Bhaskar [4] develops a refinement argument that favors this view.

support cooperative outcomes in the framework of private monitoring.

REFERENCES

[1] M. Aoyagi (1997), "Collusion in dynamic Bertrand oligopoly with correlated private signals and communication," mimeo.

[2] R. Aumann and L. Shapley (1994), "Long-term competition—a game-theoretic analysis," in N.Megiddo (Ed.), *Essays in Game Theory*, Springer-Ferlag.

[3] J. Benoit and V. Krishna (1985), "Finitely repeated games," *Econometrica* 53, pp.905-922.

[4] V. Bhaskar (2000), "The robustness of repeated game equilibria to incomplete payoff information," mimeo.

[5] V. Bhaskar and E. van Damme (1997), "Moral hazard and private monitoring," mimeo.

[6] V. Bhaskar and I. Obara (2000), "Belief based equilibria in the repeated prisoners' dilemma with private monitoring," mimeo.

[7] O. Compte (1998), "Communication in repeated games with imperfect private monitoring," *Econometrica* 66, pp.597-626.

[8] G. Ellison (1994), "Cooperation in the prisoner's dilemma with anonymous random matching," *Review of Economic Studies* 61, pp.567-588.

[9] J. Ely and J. Valimaki (1999), "A robust folk theorem for the prisoner's dilemma," mimeo.

[10] J. Friedman (1971), "A noncooperative equilibrium for supergames," *Review of Economic Studies* 38, pp.1-12.

[11] D. Fudenberg and D. Levine (1983), "Subgame-perfect equilibria of finite and infinite horizon games," *Journal of Economic Theory* 31, pp.227-256.

[12] D. Fudenberg and D. Levine (1991), "An approximate folk theorem with imperfect private monitoring," *Journal of Economic Theory* 54, pp.26-47.

[13] D. Fudenberg, D. Levine and E. Maskin (1994), "The folk theorem with imperfect public information," *Econometrica* 62, pp.997-1040.

[14] D. Fudenberg and E. Maskin (1986), "The folk theorem in repeated games with discounting or with incomplete information," *Econometrica* 54, pp.533-556.

[15] D. Fudenberg and J. Tirole (1991), *Game Theory*, MIT Press.

[16] M. Kandori and H. Matsushima (1998), "Private observation, communication and collusion," *Econometrica* 66, pp.627-652.

[17] H. Matsushima (1991), "On the theory of repeated games with private information, Part 1: Anti-folk theorem without communication," *Economics Letters* 35, pp.253-256.

[18] H. Matsushima (2000), "The folk theorem with private monitoring and uniform sustainability," mimeo.

[19] M. Piccione (1998), "The repeated prisoner's dilemma with imperfect private monitoring," mimeo.

[20] R. Radner (1986), "Repeated partnership games with imperfect monitoring and no discounting," *Review of Economic Studies* 53, pp.43-58.

[21] H. Royden (1968), *Real Analysis*, Macmillan.

[22] A. Rubinstein (1979), "Equilibrium in supergames with the overtaking criterion," *Journal of Economic Theory* 21, pp.1-9.

[23] T. Sekiguchi (1997), "Efficiency in repeated prisoner's dilemma with private monitoring," *Journal of Economic Theory* 76, pp.345-361.

[24] G. Stigler (1964), "A theory of oligopoly," *Journal of Political Economy* 72, pp.44-61.